Q&A Chat Webinar 4 (Difference in Differences II Application)

March 28th 2024

	Questions	Answers
1.	Suppose the coefficient on B1 (treat) is statistically significant i.e., treatment and controls are different at the baseline. Would this be self-selection bias? What approach does DiD use to correct the bias? Thanks!	It is not necessarily selection bias. Remember DID requires parallel trends!! Meaning that we are assuming that the change over time for the control group is what would have happened to the treatment group had the treatment group not received the transfer. Control group is our credible counterfactual of the treatment group. Not because the two groups started at the same level (i.e., equality in means at baseline) but rather they have the same trend. Now if the baseline means are very different, then it COULD be an indication that the trends are not parallel (selection bias). DID itself doesn't control for selection bias; DID assumes parallel trends.
2.	I would be happy if the data being used for these demonstrations could also be made available on the platform. This will allow learners to practice and replicate the analysis. Thanks	The data is available on the P3-Africa website.
3.	Thank you very much for your work. However, I would like to ask you if it would be possible for you to give us some exercises so that we can get a deeper understanding of what we are learning.	Please try and replicate the analysis done during the sessions. The data and DO files are on the website.
4.	Please, how can DID capture dis-adopters when evaluating impact among adopters	I assume you would have multiple time periods, then some units in the 'treated' group would switch to 'untreated'--these are the disadopters. So, the 'treat' dummy variable would become zero for these units in that time period.
5.	Is it possible for the recording to be shared between members?	Hello, yes we will upload recordings, PowerPoint slides, and the Q&A chat on our website: https://www.airp3-africa.org/applied-quantitative-analysis-webinar-series within a week's time
6.	Please where can I assess the recorded video?	The recording will be posted here: https://www.airp3-africa.org/applied-quantitative-analysis-webinar-series
7.	When using DID to estimate impact of an intervention, if in the first year some farmers accepted the intervention, but discontinue with the intervention, how can there impact be captured at the endline	See the response above to the question about 'disadopters' (Q4)
8.	You have pointed out impacts differently - i.e., after T1 and after T2. But what is the overall impact in the three-period model (baseline, T1, T2)?	What do you mean by 'overall' impact. The three period model gives impact estimates between baseline and the first follow-up, and from baseline to the second follow-up. So maybe the

		second estimate is the 'overall' impact, as it gives the impact over the entire period.
9.	"sir, is this teaching recorded? if yes, how can we have access to it"	This webinar is being recorded and will be posted to the series page here: https://www.airp3-africa.org/applied-quantitative-analysis-webinar-series
10.	What is the difference between the multiple periods DiD and Time event analysis	They are similar. In event studies, typically one models the trend in the outcome after the event in more sophisticated ways, using time series data at frequent intervals, to see the trajectory. DID isn't that refined, it is just comparing means 'before' and 'after' the event (the treatment). Event studies require much more granular (detailed) data over time.
11.	What I mean is - apart from calculating two impacts i.e., impact at T1 and impact at T2, is there a way of calculating overall impact (not period by period, but the overall)?	"See the response above to your original question. The 'overall' impact is actually the T2 interaction since it is giving the impact from baseline to endline. (Q8)
12.	What do they mean exactly when they say "fixed effect"? Is there a "random/non-fixed effect"? in that sense?	Yes, there is a random effect. WE will discuss fixed effects in a few sessions down the line!
13.	On slide 21, is it possible that there could be a convergence between the control and treatment in future time periods?	Yes, if the treatment effect dissipates or gets smaller or fades out over time. For example, training teachers in a certain method may lead to immediate gains in student achievement. But then teachers become lazy and fall back into their old habits, and we see the achievement gains go down. There are many examples of this.
14.	"in the model, DD24 and DD36 were these calculated manually? Incase I missed the code"	Yes
15.	Can we have access to the data for pratices.	You can download STATA files for today through this link: https://drive.google.com/file/d/1v1Rrhvc-6G3_cZ3hCdm88K_kifN1eCcc/view
16.	Was the data used balanced panel data? In short, were the respondents from baseline to endline the same in both control and treated groups?	Yes
17.	in regard to the first two sessions, specifically the second session "Potential outcomes and omitted variable bias II"; do we also use STATA? Are other softwares like Eviews useful?	WE tend to use STATA for everything. The first topic did not really have a 'hands-on' application.
18.	Is it possible to have access to the do file ?	You can download STATA files for today through this link: https://drive.google.com/file/d/1v1Rrhvc-6G3_cZ3hCdm88K_kifN1eCcc/view
19.	how do we interprete period24	period24 is the trend of the control group from baseline to period 24
20.	but in the reg output variable treat is not significant	What does the coefficient in front of treat measure? [The difference between the two groups at baseline. So, in fact, we hope that it is NOT significant, right?]

21.	as the treat's coeff is not significant, what can we conclude? That the treatment were not effective ?	See response above
22.	If it was unbalanced would this adversely affect the results obtained?	Only if the imbalance was systematic or selective. For example, say we lost all the women, then the estimates would be weighted more heavily by the responses/reactions of men.
23.	just for clarity, Could give an example of how to interpret this regression results.	Please see the video
24.	Could you please show us how Diff-in-diff dataset is presented on excel/stata sheet	Yes the STATA data files are on the P3-Africa website--see above for the links.
25.	Would the observation that in the second period the control group improved more than the treated group?	In Slide 7 and 8 we don't see this. However, it is possible, if the treatment effect fades out.
26.	When the areas in a country receiving treatment are not randomly selected, how can this be taken into account in difference-in-difference regression?	you could run a doubly robust diff-in-diff. If you know how treatment was assigned, you could estimate the propensity to receiving the treatment and use the propensity scores to re-weight your coefficient to account for the imbalances between treatment and comparison group.
27.	The coefficient of treat measures the DID at the baseline	Refer to slide #6. Does treat measure DID at baseline? or does it measure the differences between treatment and control at baseline? what does it mean when b1 is not significant?
28.	could Audery kindly share her do file on how she created this data set with us (me)	You can download STATA files for today through this link:https://drive.google.com/file/d/1v1Rrhvc-6G3_cZ3hCdm88K_kifN1eCcc/view
29.	How do you control for other variables that might have the effect on your outcome? (in the DiD)?	Just add them to the regression!
30.	the different of the two group before the treatment. I understand. So, I understand that there is no difference between the two groups before the intervention. thanks	!!!!!!!!!!!!!!!! :)
31.	sir, can I have access to previous webinar?	Yes, webinar materials from the previous webinar have been posted to the series page here: https://www.airp3-africa.org/applied-quantitative-analysis-webinar-series
32.	how to perform the counterfactual analysis	not sure what you mean. In these examples, the control group (for the Zambia example) is being used as the counterfactual. For the Ghana DHS example, urban is the counterfactual for rural, but that example isn't an intervention study. Rather, in that example we are just comparing trends over time among two groups.
33.	Please what is representing the treatment in Rural area? What is the intervention in the rural please?	There is no actual intervention in rural areas. Rather, we in this example we are comparing two groups: rural versus urban. So rural is analogous to the treatment group, and urban the control group (or vice-versa, it doesn't matter).
34.	I did not get the research question for this example.	The research question is whether the rural youth are improving relative to urban youth over this time period in Ghana. [AS a motivation, many African countries, including

		Ghana, are trying to increase economic opportunities for youth in rural areas so that they don't migrate to the cities.]
35.	Can WE have Access to the do file?	You can download STATA files for today through this link: https://drive.google.com/file/d/1v1Rrhvc-6G3_cZ3hCdm88K_kifN1eCcc/view
36.	how does DID work for panel data esp with staggered treatment. would you consider another session for that?	We will consider it. It is an advanced topic though, there may not be too much interest. But see the few slides we have provided, and the references.
37.	can we add an additional session on did. I am really interested in stacking data	We will consider it, thank you
38.	Is it possible to do this comparing or using two different datasets e.g. DHS vs HBS?	Only if the variables/indicators are the same.
39.	How would you then explain the coefficient on DD2008?	This is the DD estimate between 2003 and 2008. AS it isn't statistically significant, there is no difference in trends between rural and urban over this period.
40.	Please I am still not getting what makes DiD different from RCT as in RCT we still collect baseline and also some follow up. Can we use DiD method to analyze data from an RCT research given we have all it takes to do it.	Yes in fact most RCTs do still use DID to account for any small baseline differences that might occur. DID is an analytical method. RCT is a DESIGN.
41.	Faut il interpreter la DD lorsque le coeficient n'est pas statitiquement significatif? comme le cas de DD2008 = -0,387703	This is the DD estimate between 2003 and 2008. As it isn't statistically significant, there is no difference in trends between rural and urban over this period.
42.	I see the stacked data on repeated cross sections. How do I get the treatment? I didn't quite get that.	There is no actual treatment. The two groups are rural and urban, and we are comparing these two groups. WE are trying to illustrate the flexibility and wide applicability of the DID method. It isn't just for actual 'interventions', it can be used to compare any two groups over time--depending on your research question.
43.	I get you on making the data stacked. I just need to be sure - the DHS is not a panel. What does it mean to do DiD using non-panel data?	In the second example we 'stacked' the DHS for three years. DID is simply comparing changes in group means over time. You don't need a panel to do that! That is exactly the point of this example, to illustrate the power of the DID method.
44.	Can you run a DiD for more that 3 stages? If so, how would it look like?	I am prettty sure you can figure that! Add another time variable, and then (time*treat).